DOCUMENT RESUME

ED 279 695 TM 870 082

AUTHOR Rivera, Charlene
TITLE The National Assessment of Educational Progress: Issues and Concerns for the Assessment of Hispanic Students.
SPONS AGENCY Spencer Foundation, Chicago, Ill.
PUB DATE 27 Sep 86
NOTE 33p.; One of 46 papers commissioned by the Study Group on the National Assessment of Student Achievement and cited in Appendix B to their final report "The Nation's Report Card" (TM 870 049). For other papers in this group, see TM 870 050-094.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; *Educational Assessment; Educational Policy; Elementary Secondary Education; *Hispanic Americans; *Measurement Objectives; National Surveys; Racial Identification; Research Utilization; *Sampling; Spanish Speaking; Student Characteristics; *Test Construction; Testing Problems; Testing Programs
IDENTIFIERS Common Background Questionnaire; Language Minority Survey; *National Assessment of Educational Progress

ABSTRACT

The redesign of the National Assessment of Educational Progress (NAEP) has resulted in improved usefulness of the data set to conduct policy relevant research on Hispanic student achievement. Improvements include changing the procedure for ethnic identification to self-reporting, modifying sample selection, and planning for the conduct of special assessments. The Language Minority Survey (LMS) was designed to: collect data on student participation in special programs; analyze achievement of students with non-English language backgrounds; and examine the relationships between achievement and relevant school, teacher, and student attitudes. LMS, however, will not assess the progress of LM limited-English-proficient youth. Sampling procedures have incorporated grade and age-level sampling, documentation of excluded students, and Basic Incomplete Block (BIB) sampling. The "new" NAEP has been structured to address policy-relevant issues such as national concerns, human resource needs, and school effectiveness. It should now be possible, if the commitment is made, to examine the achievement of Hispanic youth, and to find ways to encourage persistence in school, factors contributing to achievement, and the affects of the effective school movement. Students' educational experiences and aspirations could be examined through the Common Background Questionnaire; however, the aspiration question does not appear in this questionnaire for the general 1986 NAEP assessment. Other suggestions for improvement are presented. (GDC)

ED279695

The National Assessment of Educational Progress:

Issues and Concerns for the Assessment of Hispanic Students

Charlene Rivera, Ed.D.
Development Associates

Paper commissioned by

THE STUDY GROUP ON THE NATIONAL ASSESSMENT OF STUDENT ACHIEVEMENT

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

C. Rivera

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 870 082

BEST COPY AVAILABLE

Revised: September 27, 1986

The National Assessment of Educational Progress:

Issues and Concerns for the Assessment of Hispanic Students

Charlene Rivera, Ed.D.
Development Associates
2924 Columbia Pike
Arlington, VA 22204

Commissioned by The Spencer Foundation
September 1986

ACKNOWLEDGEMENTS

I am grateful to many individuals who have shared valuable information about the utility of the NAEP for the study of Hispanic student achievement. Richard Duran, UC Santa Barbara, Roberto Fernandez, University of Arizona, Michael Olivas, University of Houston, Vilma Ortiz, ETS, Dorothy Waggoner, Washington, DC, Frank Alejandro, OERI, and Oscar Uribe all contributed useful information and counsel through personal or telephone interviews. Nancy Mead, Director of Operations for the NAEP, provided helpful reactions to an earlier version of this report. Richard Duran, and Oscar Uribe provided invaluable comments to an earlier draft of this report.

Introduction

The purpose of this position paper is to analyze those features in the in-school NAEP assessment that make the data base particularly useful or that limit its utility for researchers interested in using it to study factors that affect Hispanic student achievement. Because of the redesign of certain aspects of the in-school NAEP for the 1984 assessment, several important improvements were made that directly affect the utility of the NAEP data set to conduct policy relevant research. These include changing the procedure for ethnic identification, modifying the sample selection process, and planning for the conduct of special assessments. These modifications are not only timely but critical in making the data set important for the study of Hispanic youth.

In this discussion, two major issues are considered. The first relates to the assessment itself -- What are the assets and limitations of the specific components of the "new" NAEP that affect research related to Hispanic student achievement? The second consideration relates to the design of the assessment -- To what extent is the NAEP a policy relevant data base? How can the NAEP become more policy relevant for the Hispanic community?

These two issues provide the structure for the paper. Within each section specific recommendations are made relative to the improvement of the assessment as it affects the study of Hispanic youth.

Section One

Components of the NAEP Design: Study of Hispanic Student Achievement

A critical modification in the design of the 1984* NAEP Assessment that greatly enhances its attractiveness and that makes it, in fact, of use to researchers interested in studying Hispanic student achievement is the decision to use self-identification rather than visual-identification as the basis for reporting achievement. Because of its importance, this design change is the first component discussed.

Then, the focus is shifted to a consideration of the advantages and disadvantages of other components of the NAEP design -- sample selection, questionnaire development, and special assessments. While the center of interest is Hispanics, a large number of the concerns identified are of importance to researchers regardless of the race/ethnic groups being studied. The uniqueness of the discussion comes in consideration of the special factors that concern researchers interested in sorting out different components of the Hispanic population and the affect of different background variables on Hispanic student achievement.

Throughout the discussion, it is important to keep in mind that the Hispanic community and, ergo, Hispanic students identified through the NAEP are

---

* In this discussion of the in-school NAEP, the surveys are referred to by the last year in which the assessment was conducted. The assessments that are most often discussed are the three most recent surveys. They include the 1984 assessment, the 1986 general assessment, and the 1986 Language-Minority Assessment, referred to as the Language-Minority Study (LMS). References are also be made, as necessary, in the text, to earlier assessments.

not monolithic. They represent one of the fastest growing segments of the population (Chronicle, 1986).* Among the school age population, achievement continues to lag and the the drop-out statistics for this group are alarmingly high (Duran, 1983; Hispanic Policy, 1984). Thus, it is of utmost importance to be able to utilize databases such as the NAEP to conduct policy relevant research.

Race/Ethnic Identification

Hispanic student achievement was independently summarized for the first time in a 1977 report for the NAEP assessments conducted between 1971 and 1975. The report, "Hispanic Student Achievement in five Learning Areas: 1971-1975" (Crane, 1977) provided student achievement data in - Social Studies, Science, Mathematics, Career and Occupational Development, and Reading. A second special report, "Students from Homes in Which English is not the Dominant Language: Who are They and How Well do They Read?" presented the achievment of Hispanic students in the 1979 reading assessment (NAEP, 1982).

Achievement data for these reports were generally** based on a visual classification procedure where a test administrator determined a student's race/ethnicity according to the student's appearance.*** The 1984 assessment

---

* Between 1980 and 1985 Hispanics' growth rate was 16 percent compared to the population as a whole for which the growth rate was 3.3 percent.

** While not clearly stated in the 1979 report, self-identified race/ethnicity rather than visual identification was used to classify 17-year-olds and to report their achievement.

*** Prior to the 1973 assessment, race/ethnicity visual classification categories were limited to White, Black, and "Other". The "Other" category included students who did not fall into any of the other specific classifications. In 1973, the potential race/ethnic visual classification categories were expanded to include Mexican-American and Puerto Rican students. With these additions to the classification categories, the test adminstrators were advised to use a student's surname and/or speech as an aid when there was uncertainty of a student's race/ethnicity (Crane, 1977).

was the first for which race/ethnic self-identification information was collected for all age and grade levels. This was also the first assessment to report achievement on this basis rather than on appearance, although the latter identification information was also available.

The change in procedure reflects a growing acknowledgement that although flawed, self-report provides a more accurate representation of an individual's race/ethnicity (Carroll, 1981). The decision to modify the procedure was also based on mounting evidence that the visual and self-identification procedures were greatly discrepant for Hispanic students (Searls, 1977).

A study conducted by Rivera and Pennock-Roman (1986) provides a clear delineation of the issues. For this reason, the research is described, findings summarized, and recommendations highlighted.

The study grew out of the following considerations. If there are large discrepancies between visual and self-identification for some race/ethnic groups, then the validity of "who" is included in those categories is questionable. Undercounting of some subsamples may, infact, be a problem because summary characteristics reported for each subgroup, such as mean achievement, may be biased by classification errors. Furthermore, given NAEP's change in race/ethnic identification methodology, the achievement data from past and current assessments may not be comparable for some racial/ethnic categories.

Weighted data from four NAEP surveys for which both visual and self-reports* were available comprised the data source** for the study. Seven

---

* See Rivera and Pennock-Roman (1986) for a description of the race/ethnicity questions asked of students across the four assessments for which self-report information is available at any age level.

** Starting with the 1976 assessment self-report information was collected for 17-year-olds. For the 1980 assessment self-report information for both 17 and 13-year olds was collected. The 1984 assessment was the first for which self-report data was collected for all three age groups including 9-year-olds.

datasets in all were utilized, four with self-report data for 17-year-olds, two for 13 year-olds, and one for 9-year-olds. The unweighted sample sizes for the seven assessments ranged from 15,859 to 38,899. For each assessment, the concordance or agreement between the visual and the self-identification methods were compared for all NAEP subgroups -- Whites, Blacks, Hispanics, Asians, and American Indian students. While comparisons were made for all groups, it is important to keep in mind that NAEP achievement data have not and will not be reported for race/ethnic groups other than White, Black, and Hispanic students. The sample sizes for the other classifiable groups -- Asians and American Indians -- have never been, nor expected to be in the future, sufficiently large to accurately calculate achievement independently. Thus, the achievement data of these groups has been and will continue to be included in the "Other" category.

The correspondence between visual and self-report were the following. For White and Black students who represented about 71% and 13% of the sample, respectively, in any given assessment, the concordance was near perfect for Whites and 95% or better for Blacks across assessment years. The concordance rates for Asians who represented, on the average, less than 2% of the sample, ranged from 64.7% for students age 17 in 1980 to 85.9% for those age 17 in 1982. The concordance rates for American Indians who represented, on the average, less than 2% of the sample were the lowest, ranging from 7.0% for students age 9 in 1984 to 29.3% for students age 17 in 1982. For Hispanics, who constituted, on the average, 9% of the sample, the concordance rates were quite variable ranging from a low of 45.8% for 9-year-olds to a high of 74.5% for 17-year-olds in 1984.

These data confirmed that use of the visual identification procedure selects out White and Black students at about the same rate as the self-

identification procedure. However, the same cannot be said for the Hispanic, Asian, and American Indian groups. The concordance level between visual and self-identification are quite discrepant and, thus, motivate the question -- Who is or is not included in the group using either method? It should be noted that because achievement data are not reported independently for Asians or American Indians, the discrepancies do not pose the same concerns as they do for Hispanics.

Because of the wide variation in concordance rates, language use data of the Hispanic subsample was also compared for all age/grade groups sampled in the 1984 assessment. There was consistent agreement in Spanish language use and self-report of being Hispanic for students age 17 and/or in Grade 11 demonstrating that for this age group language use data provides a good validity check. However, language use data and concordance rates were not found to be as consistent for students age 13 and/or in Grade 8 or for students age 9 and/or in Grade 3.

Since the evidence regarding the validity of the self-report information of the younger age groups is dubious, Rivera and Pennock-Roman strongly recommend that a subsample of parents in a subsequent assessment be surveyed to validate race/ ethnicity. Because a parent survey of a representative group of the Hispanic subsample would provide a check on the effectiveness of the race/ ethnicity questions as they are currently asked, they recommend that such a survey be conducted across all age groups.

A parental survey will be used for a new, major educational survey, the 1988 National Educational Longitudinal Study (NELS). Parents of students sampled for the NELS will complete a questionnaire in which they will indicate their own race/ethnicity and respond to several questions related to language use in the home including language used in interacting with their child. If

such an approach were adapted for the NAEP, it would be recommended that, in addition to parental self-report of race ethnicity, a question be added in which the parent or guardian indicates the child's race/ethnicity.

If the parental validity study is not carried out for a future assessment, two other recommendations should be seriously considered as validity checks of race/ethnicity. The most easily implementable procedure would be to alter the procedure for administering the background questionnaire section where race/ethnicity and language use are elicited. The second option would be to collect two additional indicators of race/ethnicity -- place of birth and surname.

The first recommendation, that of orally administering certain parts of the Common Background Questionnaire is one that was implemented for the 1984 assessment.* Although, its utility has yet to be evaluated, this approach, it seems, would be particularly useful for the younger age groups where inconsistentcies are most prevalent.

The second recommendation could be implemented by adding a nativity question to the Common Background Questionnaire and by collecting and comparing student surnames with a standardized list of surnames coded by ethnicity. One such list is compiled and available through the U.S. Bureau of the Census; other surname lists that have been used for validity checks of self-reports and third party reports are available through state bureaus of vital statistics.** Since, visual identification information is also available, it, too, could be used to categorize students (as it was for the 1984 assessment) in cases where

---

* It involves training test administrators to verbally guide students through the race/ethnicity and language use questions, providing examples, and clarifying terms, as needed.

** For the pros and cons of using such lists see Fernandez, 1985; and, Sullivan, Gillespie, Hout, & Greeley, 1983).

classification is problematic because of conflicting responses to race/ethnicity, surname, nativity, and language use questions.

In short, race/ethnic identification represents a critical classification variable. For this reason, it is crucial to validate the procedure and to refine it as necessary. Next, issues related to sample selection are discussed.

Sample Selection

In the original design of the NAEP a deeply stratified multi-stage sampling plan was utilized. This basic design, with some procedural changes continues to be used. The stratification occurs so that in the first stage, the primary sampling units are counties selected by geographical region and community type; at the second stage, the sampling units are schools, both public and private; and at the third stage the sampling units are students within the selected schools (Moore, Chromy, & Rogers, 1974; Messick, Beaton, Lord, 1983).

With the redesign of NAEP for the 1984 assessment some modifications were made to the third stage of sampling. The changes were implemented to accomodate the introduction of several new components to the assessment. These include the incorporation of 1) grade and age-level sampling, 2) documentation of students excluded from the NAEP sample, and 3) Basic Incomplete Block (BIB) sampling. Each feature is briefly discussed with the focus on the advantages and disadvantages of these changes for the study of Hispanic student achievement.

Grade-Age Level Sampling*. The major reason for incorporating grade into the sampling plan was to make it possible to link the assessment results to school, local, and state assessment practices and, generally, to educational policy (Messick, et al, 1983, p. 23). This addition to the sampling design is particularly important for the assessment of Hispanic youth who are often overage for their grade placement. If available overtime, the NAEP data could be used as a trend barometer for grade/age level placement of Hispanic youth. It could also be used to identify the subpopulations and the characteristics of those youth who are the most at risk of being overage for their grade. In places where state assessments are conducted, age/grade level achievement comparisons can provide important policy information for use at the local level.

Excluded Student Sample. Prior to the 1984 assessment, students who were not considered to be sufficiently fluent in English or who had some other handicapping condition were excluded from taking the NAEP assessment. Traditionally, principals and/or teachers made the decision to exclude students from taking the NAEP. Until the 1984 assessment, individuals responsible for these decisions had never been required to document the reasons for the exclusion.

Thus, the introduction of a four page Excluded Student Questionnaire was an important innovation in the NAEP. For the 1986 assessment, a revised questionnaire elicits the following information for students identified as Limited English proficient (LEP)-- race/ethnicity, language background, percentage of

---

* The students sampled prior to the 1984 assessment were age 9, 13, and 17. A design change was made in the 1984 assessment so that students were and, henceforth, will be sampled by grade as well as by age. Thus, for the 1984 assessment, students were surveyed within a school if they were aged 9, 13, or 17 or if they were in grades 3, 8, or 11. For the 1986 assessment, students age 9, 13, or 17 and/or in grades 3, 7, or 11 were sampled.

LEP students in the school, percentage of time the student is provided a special language program, number of years in the program, and a rating of the students' English language proficiency. The information elicited through this questionnaire provides an understanding of the population subgroups that are not represented in the NAEP.

For the 1984 assessment 122,589 students were sampled. Of those, 6.6% (approximately 8,058) were excluded. It is important to note that approximately 20% were excluded because their English language skills were judged not to be sufficiently developed to take the NAEP reading exercises. The majority of these excluded students were from Spanish and Asian language backgrounds (Callahan, 1985).

While the documentation for excluding students is an important modification in the NAEP design, it is not a panacea. The data collected is quite limited in scope. Moreover, interpretation of who is represented in the sample could be distorted by the fact that race/ethnic identification is based on observer -- the principal or teacher -- judgement rather than on self-report.

One way to overcome the classification problem, would be to survey the students themselves or their parents or guardians. If only observer reports are collected, the parent data could, then, serve as a validity check.

BIB Sampling. As described by Messick, Beaton and Lord (1983) BIB or Basic Incomplete Block Spiraling

> ...combines the advantages of matrix sampling with those of conventional spiralling. ...it involves developing a balanced incomplete block design such that each exercise is administered the same number of times as it would be in matrix sampling, but in addition each pair of exercises is also assessed a prescribed number of times. This means that each exercise will be located in several different packages or booklets, so that many different packages must be printed for an exercise pool of a given size. The BIB spiralling of exercises also implies that many different packages, and thus

> different sets of exercises, will be administered in a particular assessment (1983, p.29).

The advantages of BIB sampling are that it makes it possible to estimate interrelationships among NAEP exercises regardless of whether they are found in the same or different booklets. Thus, "different correlations (can be) based on different student subsamples ... uniformly spread throughout the facets of the stratified sampling plan" (Messick, p.4). However, "the introduction of four subject-matter areas beginning with the 1986 assessment -- which includes reading, science, mathematics, and computer competence -- presents horrendous logistical problems if a fully balanced incomplete block design is attempted across all four areas" (Messick, p.4). For this reason a partly balanced "design was formulated that provides for ... a systematic pairing of blocks between areas to permit selected exercises in each area to be linked to exercises in each of the other three areas" (Messick, p. 5).

Because of the size of the Hispanic subsample, the advantages of BIB sampling are greatly diminished for this subgroup. Placement of an item in an assessment package that is not common to all students means that even with oversampling, sufficient numbers of Hispanic students will not be asked the same question to make it possible to link responses with other common background and subject matter areas. For Hispanics, the smallest subgroup for whom achievement is reported, the concern is to be able to link the same critical variables for all students for whom achievement is reported, be they White, Black, or Hispanic. Thus, it is important to explore how the NAEP design can be modified to take advantage of the BIB so that components of an assessment can be analyzed for all subgroups linking individual items and/or groups of items with student achievement.

Questionnaire Development

While it is acknowldeged that it is not possible to incorporate every interesting item into the Common Background Questionnaire, it is important to rethink which variables are critical for studies of special subpopulations included in the NAEP database. An example of one item that is important for the study of Hispanic student achievement is the question regarding students' educational aspirations -- Do you expect to graduate from high school? For the 1984 assessment this item was placed in a cognitive rather than a commmon exercise. Because of the high drop out rate among Hispanic youth, educational aspirations is a pivotal indicator that can be used to differentiate subsamples within the population who are at more risk than others of dropping out of school. However, because of the statistical limitations and the relatively small subsample of Hispanics, if the question of aspirations is not included in a common component of the assessment, the resulting data cannot be easily linked to student achievement.*

It is important to note that the aspiration question will be included in the 1986 LMS questionnaire. However, the question is not included in the Common Background Questionnaire administered for the general 1986 NAEP Assessment.

The background information that ideally would be added to the NAEP includes more student language background information about proficiency in English and in the primary non-English language. In addition, race/ ethnicity questions should be augmented with information about nativity, age of arrival in the United States, years of schooling outside the United States, and more detailed information about the parents' educational background and current

---

* Personal Communication with Vilma Ortiz, July 1, 1986 and with Richard Duran, August 1, 1986.

occupation. To make this information available to the NAEP three options are possible.

One choice would be to simply add the pertinent items to the Common Background Questionnaire. Another alternative would be to carefully link the race/ethnicity and language questions with those in the Census. If done properly, it should then be possible to make generalizations, based on census data, about the demographic characteristics of the students who took the NAEP. This approach could prove to be substantially useful to the NAEP and yet not essentially increase the assessment time or significantly add to the Common Background Questionnaire. Yet another option for increasing the amount of information collected about language minority students is to use a branching technique for the Common Background Questionnaire.

Using this approach, students identified to be limited English proficient or from homes where a language other than English is used would be asked to respond to a special series of questions about nativity, years of schooling in the United States and in the students' homeland, non English language use, and English language proficiency. In addition, they would respond to a majority of the the general questions responded to by the rest of the population. This means that some general questions for these students would be shortened or not asked. However, if this option is used, it will be exceedingly important to make certain that there is sufficient commonality in the questions asked across all subgroups to ensure comparability while maintaining the integrity of the questions that have traditionally been used in the NAEP.

It should be noted that the branching approach has been successfully used for the 1986 NAEP Adult Assessment. The branching, in this case, occurred orally. That is interviewers branched or asked different questions based on a respondents' responses. The technique will also be used for the 1988 NELS

assessment of 8th graders. In this instance, the respondents will be reacting to written instructions. These experiences should provide useful evaluative information, if it is decided to try the technique in the NAEP.

Regardless of the approach adopted, it will be important for NAEP staff to work closely with educators and researchers aware of the relevant issues facing Hispanic youth. Priorities will then need to be made selecting the issues to be incorporated into the common components of the NAEP.

Special Assessments

The Language Minority Survey (LMS) represents one of several special studies that have been proposed as a result of the redesign of the NAEP. To date, it is the only one that has actually been funded and scheduled. Its purpose is to supplement the educational achievement information available in the NAEP data base for language minority youth. Specifically, the goals for the LMS are: 1) to provide program information about the participation of students in special programs; 2) to analyze the academic progress of students who come from non English language backgrounds; and 3) to examine the relationships between achievement and relevant school, teacher, and student attitudes (Duran and Baratz, 1984).

This special assessment is particularly important because it will provide one of the few databases with achievement information on the Hispanic language minority (LM*) population. It is important to note that this assessment is limited in that it will not provide for the collection of information about the

---

* LM refers to students in whose home a non-English language is typically spoken. Such students may include those who speak only English; those whose English is fluent enough to benefit from instruction in academic subjects offered in English, as well as those whose English proficiency is limited.

This definition is found in the 1986 special Language-Minority Teacher and School Characteristics and Policies Questionnaires.

progress of language minority limited-English-proficient (LM-LEP*) youth who are most often excluded from taking the general NAEP assessment.

A major difference between the 1986 general NAEP survey and the Language Minority Survey needs to be pointed out. While the sampling for both assessments is based on a nationally representative sample of students, the LMS is being conducted, for the most part, in a separate set of 240 schools with high concentrations of language minority students. A small number of students who take the main NAEP will also be selected for the LMS study. However, the number of language minority students identified through the main NAEP are limited. Thus, to ensure a large enough group to make analyses possible, oversampling is necessary.

Because the groundwork has been done (i.e. questionnaire items have been developed and tested and a sampling design developed), it is strongly recommended that it either be replicated periodically or that key portions of the assessment be incorporated into the general NAEP assessments.

As part of the regular NAEP assessments, the LMS could be administered on a regular cycle, perhaps every four to five years. The ideal situation would be to conduct the assessments at intervals that would make it possible to assess the same students at different points in time. By devising techniques to track students, in addition to cross-sectional data analysis, it would make

---

* LM-LEP refers to students whose dependence on a language other than English interferes with his or her progress in school. This includes students with no proficiency in English. LM-LEP students are a subset of LM students.

This definition is found in the 1986 special Language-Minority Teacher and School Characteristics and Policies Questionnaires.

longitudinal analysis feasible. An added advantage of this option is that extensive and specialized information would be collected for the language minority populations represented in the NAEP. Moreover, this approach would make it possible to examine the data over time and to link the data with other national data sets.

While this is a reasonable approach, it is not without its limitations. In order to avoid distortions about who is language minority, it is important to be able to make linkages with the general NAEP. This means that, in addition to administering common assessment components, certain questions used to identify subgroups of the population would need to be incorporated into the NAEP assessments. This would make comparisons across the two assessments more feasible and would provide a broader picture of where the students come from and what their schools and teachers are like.

The second option of only incorporating key portions of the LMS survey into the regular NAEP is a very viable, easily implemetable alternative that involves the careful selection of critical components of the LMS. Not only will it be important to select questionnaire items, but also to carefully consider issues related to sampling and implementation. If this option is selected, it will be important for the NAEP staff to confer with individuals who are skilled in the demography and language background of the language minority populations to be studied. For Hispanics, it will be particularly important for the sampling to be broad, yet representative since they are the only language minority subgroup for whom achievement data can reliably be independently reported.

Section Two

NAEP: Policy Relevancy

The goal for this second section of the paper is to highlight how the "new" NAEP can become even more ive as a policy tool, particularly as it relates to the study of Hispanic youth. In particular, the concern will be on identifying those areas in the NAEP that affect its utility as a policy tool and on making recommendations toward that end.

The architects who conceived the the NAEP responded to the political and social realities of the times. An important consideration in the late 1960's was with how NAEP results were to be used and reported. In particular, there were three concerns: 1) that educational programs not be controlled by the federal government; 2) that assessment results not be reportable by state; and 3) that the federal government not establish national performance standards or a national curricula based on the results of the assessment.

In contrast, the designers of the "new" NAEP integrated innovations into the assessment to improve its "interpretability" and "policy relevance" (Messick, et al, p. 11). The "new" NAEP is structured to address issues relative to "student competencies as they relate to national concerns; student achievement and attitudes as they relate to human resource needs; and student achievement as it relates to school iveness" (Messick, et al, 1983, p. 11). It is intended to be used as a vehicle for not only providing information about the current status of the educational achievement of youth in the United States but to help identify areas in the educational process that are and/or that will be critical determinants of future academic success. Thus, the goal is to not only make NAEP findings relevant at the national level but to also make them useful at the state and local levels. Because policy relevancy was not a goal

for the earlier assessments and because of the visual classification criteria used to report achievement, the pre-1984 assessments are not, in fact, useful to conduct policy relevant research of Hispanics.

While the designers of other national data sets such as the Census have invited Hispanic scholars to assist in making the data sets more relevant to the Hispanic community, this has not been a priority for the NAEP nor is it a goal at this point in time. Nevertheless, the 1986 LMS does represent a significant attempt to collect relevant information about language minority students. However, since it is not an established ongoing innovation, at this point, the LMS must be viewed tentatively as a promising first step.

Because their academic achievement continues to lag below that of the population as a whole (Duran, 1983), factors that affect Hispanic students' success in school are of critical importance to educators and policymakers at all levels. While many questions can be asked about the educational achievement of Hispanics, four frame the policy issues that are of critical concern to policymakers. Posed in the form of questions, they are:

- Why does the achievement of Hispanic youth at the elementary and secondary levels lag behind that of the nation as a whole?
- What are the factors that contribute to greater success of Hispanic youth in elementary and secondary school?
- What can be done at the elementary and secondary school level to improve the educational experience of Hispanic youth so as to keep them from dropping out and to encourage them to continue onto postsecondary education?
- What, if any, have been the affects of the effective school movement on the achievement of Hispanic school age youth?

Because of the redesign of the NAEP, the potential exists to address these issues at a level that has not been previously possible. Of particular note, is the feasibility of not only linking student background characteristics but

teacher and school policies characteristics to student achievement. Moreover, for Hispanics, the use of self rather than visual identification to report achievement is an important improvement to the quality of the NAEP data. However, as noted earlier, the advantage of the greatly expanded data base is seriously diminished by the problems associated with the BIB design.

Specifically, the 1984 NAEP assessment provides information about:

- students excluded from the assessment;
- student demographic characteristics including information about English and non English language use in the home;
- educational experiences which includes information relative to curricular offerings, school educational policies, school resources, etc.;
- teacher background including areas of certification, professional training, and current assignment; and,
- institutional characteristics including school socioeconomic status, special subject area course offerings, school integration, and special services provided.

The importance of having this range of data available in the NAEP is that each factor or combination of factors can be theoretically linked to student achievement. While the data available in the "new" NAEP are far from perfect, they provide a basic starting point for exploring issues that affect Hispanic student achievement. Moreover, because of the improved quantity and quality of the data now collected, research conducted using the NAEP can validly be used to make limited generalizations about the Hispanic subsample.

However, if the NAEP is to provide policy relevant data for Hispanic youth beyond that available as a result of the 1984 and 1986 assessments, a longer range plan will need to be developed. In projecting toward this end, several areas of the current approach are considered in Section Two. The issues discussed parallel those in Section One. They include consideration of the

procedure for identifying race/ethnicity in the NAEP, sample selection, questionnaire development, and the Language Minority Survey. In addition, issues related to trend analysis are discussed.

Classification Variables for Race/Ethnicity

By improving the validity of the race/ethnicity information and by validating the language use data, it will be possible to generalize more confidently about whom the NAEP represents. Such conviction is a qualitative component that is indispensible for making the NAEP attractive to researchers interested in conducting research relative to Hispanic student achievement.

To ensure that race/ethnicity variables continue to be improved, it will be important for the NAEP staff to encourage researchers to conduct studies to assess changes in race/ethnicity identification procedures and to evaluate the quality of new components that might later be added to the data base. For this reason it will be important to maintain links with researchers conducting studies focused on Hispanics using NAEP, to prompt feedback on the quality of the race/ethnicity data, and to elicit suggestions for improving the quality of this critical component of the survey.

Sample Selection

For the LMS, provision has been made to oversample Hispanic students. However, a committment to oversample Hispanics for future general NAEP assessments has not been made. This is a critical issue because the NAEP will be relevant to the Hispanic community, only if this innovation in the

design is made, making the data suitable for investigation by researchers and important from a policy perspective.

The policy relevancy of three other components of the sampling design discussed earlier -- grade/age level sampling, sampling of students excluded from the NAEP, and BIB sampling -- are also again briefly considerd. Together with oversampling, these innovations contribute to the quality of the NAEP data set.

Grade/age level sampling is an improvement to the NAEP design that heightens the quality of the data and in particular its policy relevance for Hispanic youth. The ability to compare grade/age achievement is a particular advantage because it makes it possible to analyze grade delay across all age groups. For Hispanics, this is a critical factor that greatly affects educational achievement.

The introduction of the excluded student sampling now makes it possible to know who is not included in the assessment and the reasons for their exclusion. While this is an important development in the NAEP and an important source of information about the Hispanic subsample, it is not without its limitations.

As pointed out earlier, the BIB design, as it is currently operationalized in the NAEP has also proven to be problematic for the interpretation of certain parts of the data for the Hispanic sample. Because of the limited size of the Hispanic sample in the NAEP, statistical limitations make it difficult to link certain important background variables to achievement. While it is not easy to make suggestions about how to resolve the problems associated with the BIB design, it is an important enough issue to recommend that the background information included in the BIB be critically evaluated in the light of its interpretability and generalizability for the Hispanic subsample. From

a policy perspective, a resolution to the problematic issues associated with the BIB design for the Hispanic sample is essential.

Overall, if NAEP is to be relevant to the Hispanic community, greater effort will need to be made in planning how to validate the information collected on the excluded student questionnaire. In addition, it will be important to explore options for overcoming the problems associated with BIB spiralling. Most importantly, if the NAEP is to be of use to the Hispanic community in the future, a committment will need to be made to oversample Hispanics on a regular basis.

Questionnaire Development

The questionnaires developed for the general NAEP take into consideration variables that are of general interest. The individual items were selected by the NAEP staff because they provide continuity across assessments and because individually or in combination, when analyzed, will yield policy relevant data.

Three topics that would greatly increase the policy relevance of the NAEP for Hispanics -- nativity, expanded language use data, and educational background data -- were discussed in Section One. They represent key classification variables that are critical and that will, undoubtedly, influence the extent to which the data base will be used to study Hispanic student achievement. The inclusion of these classification variables or alternatives that link the NAEP data sets with the Census will make it extremely attractive to researchers. Moreover, the continuous oversampling of Hispanic youth has the promise of providing a greatly enhanced data set to conduct policy studies at a variety of levels.

In order to evaluate the policy relevance of the items currently included in the the NAEP questionnaires, it is recommended that a representative group

of demographers, social science researchers, and policymakers who are knowledegable about the factors that affect Hispanic student achievement be identified. This group should become involved in a review of currently used questionaire items and provide input into the development of revised or new items. Communication should also be maintained with researchers developing questionnaire items for other relevant data bases such as the NELS. In sum, because questionnaire development represents a critical foundation for the NAEP data collection, it will be important to devise structures that will allow for both formal and informal feedback from individuals sensitive to and knowledgeable about the Hispanic community.

Language Minority Survey

The committment to seek funding for special assessments, in particular the study of limited English speaking students, is important to note. However, as previously discussed, the 1986 assessment is planned as a one-time effort. Since it represents an important data source about Hispanic youth, it would be worthwhile to assess the feasibility of replicating the LMS on a regular cycle. At the very least, the special instruments developed for the assessment and the insights gained from its implementation should be evaluated so that crucial aspects of the assessment can be incorporated into the general NAEP, a low cost alternative. The additional data collected as a result would greatly increase the attractiveness of the data base for researchers interested in studying Hispanic student achievement. This latter option, to some extent, is even more appealing than the former. In addition to being cost efficient, incorporating relevant critical components from the LMS into the general NAEP would mean greater comparability across all subgroup samples. From a policy perspective,

it would provide a data base that would be greatly enhanced and that could be used to conduct policy studies at a variety of levels.

Reporting Trend Data

Trend reports compare achievement across assessments over time. The latest reading report, for example, compares trends from 1971 to 1984 and the writing report covers writing assessments from 1974 to 1984 (ETS, 1985; Applebee, Langer, and Mullis, 1986). It is important to note that the classification system for reporting data for the trend reports is based on visual identification. The reason for this is that visual identification data provide the only consistent classification scheme across assessment years for all age and grade levels.

Since the use of visual classification has been shown to be problematic for Hispanics, this approach needs to be reconsidered (Rivera & Pennock-Roman, 1986). The trend reports, based on visual rather than self report data, may not, in fact, be providing an accurate picture of Hispanic student achievement over time. Evidence to this effect has been found in a study currently being conducted by Pennock-Roman and Rivera for which reading trend data is being examined for Hispanics classified by both the visual and self-identification methods (in progress).

Because it is important not to have a distorted picture of Hispanic student achievement, it is strongly recommended that, henceforth, trend reports either present findings for Hispanics based on both visual and self-identification methods or solely on self-identification. Although both methods have flaws, the advantages of self-report outweigh its limitations. For this reason, it is recommended that in the future trend reports be based on a self-identification classification. The approach would also provide

consistency with the way data has been reported for the 1984 assessment and will, henceforth, be reported for future assessments. In the trend analysis this might mean introducing Hispanics at a later point in time. Or it could require a separate trend analysis beginning at the points in time when self report data become available for Hispanic students. Whatever the approach, the important issue is that every effort should be made to provide trend reports that are based on self rather than on visual classification.

Conclusion

The topics discussed in Sections One and Two of this paper provide a perspective on the important design issues for the NAEP. These include a consideration of components of:

- the validation of race/ethnicity data;
- sampling issues;
- questionnaire development;
- reporting of trend data; and
- the Language Minority Survey.

For each topic, specific recommendations were made relative to the improvement of the NAEP design and its policy relevancy for the study of Hispanic youth.

Because the quality of the data collected through the NAEP is limited, it is critical to have the advice of expert scholars and practitioners interested in issues related to Hispanic student achievement. In particular, it is important to document that the data that has been collected for Hispanic youth is not only valid but policy relevant. More importantly, if a committment exists to make the NAEP responsive to the Hispanic community, it will be crucial that planning for future assessments take into account the accumulated knowledge gained as a result of the 1984 and 1986 assessments. This

information, together with the goals for any new assessments should guide the developers in creating a responsive research design that will meet the needs of the NAEP generally and of the Hispanic community in particular.

Assuming that the committment is made, it is recommended that in addition to inviting scholars who are experts on Hispanic student achievement to periodic committee meetings, a standing committee of scholars with expertise in sampling design, demographics, and social science research be invited to provide input into the planning of future NAEP assessments. In particular, it will be important for such a group to contribute to the plan for data analysis. The role of such a group, which could include as few as three people or as many as ten or twenty individuals, would be that of a resource and a sounding board to the NAEP staff. In addition, during the planning stages of new assessments, they could also serve as delegates of the NAEP to the community of scholars and educators interested in using the NAEP to conduct research focused on the achievement of Hispanic youth.

Further, to increase interest and participation of researchers generally in the NAEP, it is recommended that a concerted effort be made to inform educators and scholars at all levels, particularly graduate students of the existence and purpose of the NAEP, the status of the research conducted using the data set, and the availability of the data to conduct research. The Visiting Scholar Program, established by the NAEP in 1983, should also continue to promote opportunities for the study of minority groups' student achievement.

Wirtz and Lapointe (1982) point out, that the NAEP will not have reached its full potential until it becomes an instrument for formulating effective educational policy. The NAEP, as it is currently designed, is a <u>potentially</u> powerful policy tool that makes it especially attractive for the study of

Hispanic youth. Because the NAEP is the only database that provides academic achievement information for youth below the tenth grade, it represents an important data source for analyzing the relationship between and/or among student, teacher, and school background variables and achievement at a variety of levels. As such, its potential to provide relevant data into what constitute effective schooling situations for Hispanic youth should be actualized. Moreover, because the NAEP design is dynamic, it is imperative that researchers and educators, concerned about the factors that affect the achievment of Hispanic youth, their schools, and the effects of school reforms on this group, have the opportunity to provide input to the designers and implementors of the NAEP.

REFERENCES

Applebee, Arthur N., Langer, Judith A., Mullis, Ina V. (1986). NAEP Writing Trends Across the Decade, 1974-84. Prepublication Copy. Princeton, NJ: Educational Testing Service.

Callahan, Jill A. (1985). 1983-1984 NAEP Reading and writing exercises: Data collection and scoring procedures. Princeton, NJ: Educational Testing Service.

Carroll, John J. (Ed.). (1981, January). Research to improve the quality of race/ethnic data collected by the federal government. Collection of papers given at the Annual Meeting of the American Statistical Association in Houston, TX. Washington, DC: U.S. Department of Health and Human Services Social Security Administration Office of Policy Office of Research and Statistics.

Census Bureau Report on Hispanic Populations in the United States. (1986, February). Chronicle of higher education.

Crane, Robert. Hispanic student achievement in five learning areas: 1971-75. Denver, CO: Education Commission of the States.

Duran, Richard. (1983). Hispanics' education and background. New York: College Examination Board.

Fernandez, Edward W. (1975). Comparison of surname and persons of Spanish origin in the United States. Technical paper No 38. Washington, DC: Government Printing Office.

Messick, Samuel, Beaton, Albert, & Lord, Frederick. (1983). National assessment of educational progress reconsidered: A new design for a new era. Princeton, NJ: Educational Testing Service.

Moore, R. Paul, Chromy, James R., Rogers, W. Todd. (1974). The national assessment approach to sampling. Denver, CO: Education Commission of the States.

NAEP. (1982). Students from homes in which English is not the dominant language: Who are they and how well do they read?. Denver, CO: Education Commission of the States.

NAEP. (1986). The reading report card: Progress toward excellence in our schools. Princeton, NJ: Educational Testing Service.

Pennock-Roman, Maria & Rivera, Charlene. (1986, April). Ethnic identification procedures and reports of mean reading achievement in two NAEP assessments. Paper presented at the Annual Meeting of the American Education Research Association.

Rivera, Charlene & Pennock-Roman, Maria. (1986). M. A Comparison of Race/Ethnicity Identification Methods in the National Assessment of Educational Progress. Research Report. Princeton, NJ: Educational Testing Service.

Searls, D. (1977, July). Background Variables. ANAC Minutes. Denver, CO: Education Commission of the States.

Sullivan, Teresa A., Gillespie, Francis P., Hout, Michael, & Greeley, Andrew W. Surname versus self-identification in the analysis of Hispanic data. Proceedings of the American Statistical Association.

Wirtz, Willard & Lapointe, Archie. (1982). Measuring the quality of education: A report on assessing educational progress. Washington, DC: Carnegie Corportation, Ford Foundation, Spencer Foundation.